

MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION

CLAIM OF PRIORITY

[0001] The present application claims priority to U.S. Provisional App. No. 61/676,883, filed on July 27, 2012, which is incorporated by reference herein in its entirety.

CROSS REFERENCE TO RELATED APPLICATIONS

[0002] The present application is related to commonly-owned U.S. Pat. App. No. 12/504,131, filed on July 16, 2009, now U.S. Pat. No. 8,417,870, U.S. Pat. App. No. 12/761,179, filed on April 15, 2010, U.S. Pat. App. No. 13/287,042, filed on November 1, 2011, and U.S. Pat. App. No. 13/287,081, filed on November 1, 2011, each of which is incorporated herein by reference in its entirety.

FIELD

[0003] The disclosure herein is related generally to memory modules, and more particularly to multi-rank memory modules and methods of operation.

BACKGROUND

[0004] With recent advancement of information technology and widespread use of the Internet to store and process information, more and more demands are placed on the acquisition, processing, storage and dissemination of vocal, pictorial, textual and numerical information by microelectronics-based combination of computing and communication means. In a typical computer or server system, memory modules are used to store data or information. A memory module usually includes multiple memory devices, such as dynamic random access memory devices (DRAM) or synchronous dynamic random access

memory devices (SDRAM), packaged individually or in groups, and/or mounted on a printed circuit board (PCB). A processor or a memory controller accesses the memory module via a memory bus, which, for a single-in-line memory module (SIMM), can have a 32-bit wide data path, or for a dual-in-line memory module (DIMM), can have a 64-bit wide data path.

[0005] The memory devices of a memory module are generally organized in ranks, with each rank of memory devices generally having a bit width. For example, a memory module in which each rank of the memory module is 64 bits wide is described as having an "x64" or "by 64" organization. Similarly, a memory module having 72-bit-wide ranks is described as having an "x72" or "by 72" organization.

[0006] The memory capacity or memory density of a memory module increases with the number of memory devices on the memory module. The number of memory devices of a memory module can be increased by increasing the number of memory devices per rank or by increasing the number of ranks.

[0007] In certain conventional memory modules, the ranks are selected or activated by control signals from a processor or memory controller during operation. Examples of such control signals include, but are not limited to, rank-select signals, also called chip-select signals. Most computer and server systems support a limited number of ranks per memory module, which limits the memory density of the memory modules that can be used in these computer and server systems.

[0008] For memory devices in such as a memory module to be properly accessed, distribution of control signals and a control clock signal in the memory module is subject to strict constraints. In some conventional memory modules, control wires are routed so there is an equal length to each memory component, in order to eliminate variation of the timing of the control signals and the control clock signal between different memory devices in the memory modules. The balancing of the length of the wires to each memory devices compromises system performance, limits the number of memory devices, and complicates

their connections.

[0009] In some conventional memory systems, the memory controllers include leveling mechanisms for write and/or read operations to compensate for unbalanced wire lengths and memory device loading on the memory module. As memory operating speed and memory density continue to increase, however, such leveling mechanisms are also insufficient to insure proper timing of the control and/or data signals received and/or transmitted by the memory modules.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] FIG. 1 is a diagram illustrating a memory system including at least one memory module according to one embodiment.

[0011] FIGS. 2A-2D are each a diagrams illustrating interactions among components in a a memory module according to certain embodiments.

[0012] FIG. 3 is a diagram illustrating one of a plurality of data buffers in a memory module according to one embodiment.

[0013] FIGS. 4A-4B are each a diagram illustrating data and data strobe signal lines coupled to memory devices in a memory module according to certain embodiments.

[0014] FIGS. 5A-5B are diagrams illustrating different numbers of memory devices that can be coupled to each data buffer in a memory module according to certain embodiments.

[0015] FIG. 6 is a diagram illustrating a control circuit in a data buffer according to certain embodiments.

[0016] FIG. 7 is a diagram illustrating control signals from a module control device to a plurality of data buffers in a memory module according to certain embodiments.

[0017] FIG. 8 is a timing diagram illustrating alignment of module control signals with respect to module clock signals.

[0018] FIG. 9 is a diagram illustrating a metastability detection circuit and signal adjustment circuit in a data buffer according to certain embodiments.

[0019] FIGS. 10A-10C are diagrams illustrating a metastability detection circuit according to certain embodiments.

[0020] FIGS. 10D is a diagram illustrating a signal adjustment circuit according to certain embodiments.

[0021] FIGS. 11A-11B are diagrams illustrating a metastability detection circuit and signal adjustment circuit, respectively, according to certain embodiments.

[0022] FIGS. 12A-12B are a timing diagrams illustrating a write operation and a read operation, respectively, performed by a memory module according to one embodiment.

[0023] FIG. 13 is a diagram illustrating a delay control circuit in a data buffer according to certain embodiments.

[0024] FIG. 14 is a diagram illustrating a DQ or DQS routing circuit in a data buffer according to an embodiment.

[0025] FIG. 15 a diagram illustrating a DQS routing circuit having a delay circuit in a data buffer according to an embodiment.

[0026] FIG. 16 a diagram illustrating a DQ routing circuit having a delay circuit in a data buffer according to an embodiment.

[0027] FIG. 17 is a diagram illustrating a delay circuit in a DQ or DQS routing circuit according to an embodiment.

[0028] FIG. 18 is a flowchart illustrating a method for data edge alignment according to embodiments.

[0029] FIG. 19 is a diagram illustrating a control circuit in a data buffer according to certain embodiments.

DESCRIPTION OF EMBODIMENTS

[0030] A memory module according to one embodiment includes memory devices organized in groups, a module control device, and data buffers (DB). The data buffers are sometimes referred to herein as buffer circuits, isolation devices (I.D.) or load reduction devices. The memory module is operable to perform memory operations in response to memory commands (e.g., read, write, refresh, precharge, etc.), each of which is represented by a set of control/address (C/A) signals transmitted by the memory controller to the memory module. The C/A signals may include, for example, a row address strobe signal (/RAS), a column address strobe signal (/CAS), a write enable signal (/WE), an output enable signal (/OE), one or more chip select signals, row/column address signals, and bank address signals. The memory controller may also transmit a system clock signal to the memory module. In one embodiment, the C/A signals and the system clock signal are received by the module control device, which generates a set of module command signals and a set of module control signals in response to each memory command from the memory controller. The module command signals are transmitted by the module control device to the memory devices via module C/A signal lines, and the module control signals (referred sometimes herein as module control signals) are transmitted by the module control device to the buffer circuits via module control signal lines.

[0031] The buffer circuits are associated with respective groups of memory devices and are distributed across the memory module at positions corresponding to the respective groups of memory devices. Thus, during certain high speed operations, each module control signal may arrive at different buffer circuits at different points of time across more than one clock cycle of the system clock. Also, each buffer circuit associated with a respective group of memory devices is in the data paths between the respective group of memory devices and the memory controller. Thus, the memory controller does not have direct control of the memory devices. In one embodiment, each group of memory devices include at least two subgroups, each subgroup including at least one memory device. Each buffer circuit is configured to select a subgroup in the respective group of memory devices to communicate data with the memory controller in response to the module control signals.

Thus, the memory module can have more ranks of memory devices than what is supported by the memory controller.

[0032] In one embodiment, each buffer circuit includes metastability detection circuits to detect metastability condition in the module control signals and signal adjustment circuits to adjust the module control signals and/or a module clock signal to mitigate any metastability condition in the module control signals.

[0033] Further, in one embodiment, each buffer circuit includes signal alignment circuits that determine, during a write operation, a time interval between a time when one or more module control signals are received from the module control circuit and a time when a strobe or data signal is received from the memory controller. This time interval is used during a subsequent read operation to time transmission of read data to the memory controller, such that the read data arrives at the memory controller within a time limit in accordance with a read latency parameter associated with the memory system.

[0034] FIG. 1 shows a system 100 including a memory controller (MCH) 101 and one or more memory modules 110 coupled to the MCH by a memory bus 105, according to one embodiment. As shown, the memory bus includes C/A signal lines 120 and groups of system data/strobe signal lines 130. Also as shown, each memory module 110 has a plurality of memory devices 112 organized in a plurality of ranks 114. Each memory module 110 further includes a module control circuit (module controller or module control device) 116 coupled to the MCH 101 via the C/A signal lines 120, and a plurality of buffer circuits or isolation devices 118 coupled to the MCH 101 via respective groups of system data/strobe signal lines 130. In one embodiment, the memory devices 112, the module control circuit 116 and the isolation devices 118 can be mounted on a same side or different sides of a printed circuit board (module board) 119.

[0035] In the context of the present description, a rank refers to a set of memory devices that are selectable by a same chip select signal from the memory controller. The number of ranks of memory devices in a memory module 110 may vary. For example, as shown, each

memory module 110 may include four ranks of memory devices 112. In another embodiment, the memory module 110 may include 2 ranks of memory devices. In yet another embodiment, the memory module may include six or more ranks of memory devices 112.

[0036] In the context of the present description, a memory controller refers to any device capable of sending instructions or commands, or otherwise controlling the memory devices 112. Additionally, in the context of the present description, a memory bus refers to any component, connection, or groups of components and/or connections, used to provide electrical communication between a memory module and a memory controller. For example, in various embodiments, the memory bus 105 may include printed circuit board (PCB) transmission lines, module connectors, component packages, sockets, and/or any other components or connections that provide connections for signal transmission.

[0037] Furthermore, the memory devices 112 may include any type of memory devices. For example, in one embodiment, the memory devices 112 may include dynamic random access memory (DRAM) devices. Additionally, in one embodiment, each memory module 110 may include a dual in-line memory module (DIMM).

[0038] Referring to FIG. 2A, which illustrates one memory module 110 according to an embodiment, the module control device 116 receives system memory commands represented by a set of system control/address (C/A) signals from the MCH 101 via signal lines 120 and generates module command signals and module control signals based on memory commands from the system. The module control device 116 also received a system clock MCK and generates a module clock signal CK in response to the system clock signal MCK. The MCK signal may include a pair of complementary clock signals, MCK and $\overline{\text{MCK}}$, and the module clock signal may include a pair of complementary clock signals CK and $\overline{\text{CK}}$.

[0039] Examples of the system C/A signals include, but are not limited to, Chip Select (or /CS) signal, which is used to select a rank of memory devices to be accessed during a

memory (read or write) operation; Row Address Strobe (or /RAS) signal, which is used mostly to latch a row address and to initiate a memory cycle; Column Address Strobe (or /CAS) signal, which is used mostly to latch a column address and to initiate a read or write operation; address signals, including bank address signals and row/column address signals, which are used to select a memory location on a memory device or chip; Write Enable (or /WE) signal, which is used to specify a read operation or a write operation, Output Enable (or /OE) signal, which is used to prevent data from appearing at the output until needed during a read operation, and the system clock signal MCK.

[0040] Examples of module command signals include, but are not limited to module /CS signals, which can be derived from the system /CS signals and one or more other system C/A signals, such as one or more bank address signals and/or one or more row/column address signals; a module /RAS signal, which can be, for example, a registered version of the system /RAS signal; a module /CAS signal, which can be, for example, a registered version of the system /CAS signal; module address signals, which can be, for example, registered versions of some or all of the address signals; a module /WE signal, which can be, for example, a registered version of the system /WE signal; a module /OE signal, which can be, for example a registered version of the system /OE signal. In certain embodiments, the module command signals may also include the module clock signal CK.

[0041] Examples of module control signals include, but are not limited to a mode signal (MODE), which specifies a mode of operation (e.g., test mode or operating mode) for the isolation devices 118; one or more enable signals, which are used by an isolation device to select one or more subgroups of memory devices to communicate data with the memory controller; and one or more ODT signals, which are used by the isolation devices to set up on-die termination for the data/strobe signals. In one embodiment, the module control signals are transmitted to the isolation devices 118 via respective module control signal lines 230. Alternatively, the module control signals can be packetized before being transmitted to the isolation devices 118 via the module control signal lines and decoded/processed at the isolation devices.

[0042] Module control device 116 transmits the module command signals to the memory devices 112 via module C/A signal lines 220. The memory devices 112 operate in response to the module command signals to receive write data or output read data as if the module command signals were from a memory controller. The module control device transmits the module control signals together with the module clock signal CK to the isolation devices 118 via module control signal lines 230. As shown in FIG. 2, at least some of the memory devices in a same rank share a same set of module C/A signal lines 220, and at least some of the isolation devices 118 share a same set of module control signal lines 230.

[0043] As shown in FIGS. 2A and 2B, each rank 114 includes N memory devices, where N is an integer larger than one. For example, a first rank includes memory devices $M_{11}, \dots, M_{i1}, M_{i+1,1}, \dots, M_N$, a second rank includes memory devices $M_{12}, \dots, M_{i2}, M_{i+1,2}, \dots, M_{N,2}$, and so on. In one embodiment, the memory devices 112 are also organized in groups or sets, with each group corresponding to a respective group of system data/strobe signal lines 130 and including at least one memory device from each rank. For example, memory devices M_{11}, M_{12}, M_{13} , and M_{14} form a first group of memory devices, memory devices M_{i1}, M_{i2}, M_{i3} , and M_{i4} form an i^{th} group of memory devices, and so on.

[0044] As shown, the isolation devices 118 are associated with respective groups of memory devices and are coupled between respective groups of system data/strobe signal lines 130 and the respective groups of memory devices. For example, isolation device ID-1 among the isolation devices 118 is associated with the first group of memory devices M_{11}, M_{12}, M_{13} , and M_{14} and is coupled between the group of system data/strobe signal lines 130-1 and the first group of memory devices, isolation devices ID- i among the isolation devices 118 is associated with the i^{th} group of memory devices M_{i1}, M_{i2}, M_{i3} , and M_{i4} and is coupled between the group of system data/strobe signal lines 130- i and the i^{th} group of memory devices, and so on.

[0045] In one embodiment, each group or sets of memory devices are coupled to the associated isolation device 118 via a set of module data/strobe lines 210. Each group or set of memory devices is organized in subgroups or subsets, with each subgroup or subset

including at least one memory device. The subgroups in a group of memory devices may be coupled to the associated isolation device 118 via a same set of module data/strobe lines 210 (as shown in FIG. 2A) or via respective subsets of module data/strobe lines 210 (as shown in FIG. 2B). For example, as shown in FIG. 2B, in the first group of memory devices, memory devices M_{11} and/or M_{13} form a first subgroup, and memory devices M_{12} and/or M_{14} form a second subgroup; in the i^{th} group of memory devices, memory devices M_{i1} and/or M_{i3} form a first subgroup, and memory devices M_{i2} and/or M_{i4} form a second subgroup; and so on. The first subgroup of at least one memory device in each group of memory devices is coupled to the associated isolation device 118 via an associated first subset of module data/strobe lines YA, and the second subgroup of at least one memory device in each group of memory devices is coupled to the associated isolation device via an associated second subset of module data/strobe lines YB, as shown. For example, memory devices M_{11} and/or M_{13} form the first subgroup are/is coupled to the isolation device ID-1 via the corresponding first subset of module data/strobe lines YA-1, and memory devices M_{12} and/or M_{14} form the second subgroup are/is coupled to the isolation device ID-1 via the corresponding second subset of module data/strobe lines YA-2.

[0046] In one embodiment, the isolation devices 118 are in the data paths between the MCH 101 and the memory module 110 and include data buffers between the MCH 101 and the respective groups of memory devices. In one embodiment, each isolation device 118 is configured to select a subgroup in the respective group of memory devices to communicate data with the MCH 101 in response to the module control signals, such that the memory module can include more ranks than what is supported by the MCH 101. Further, each isolation devices 118 is configured to isolate unselected subgroup(s) of memory devices from the MCH 101 during write operations, so that the MCH sees a load on each data line that is less than a load associated with the respective group of memory devices. In one embodiment, the MCH sees only a load associated with one memory device on each data/strobe signal line during write operations.

[0047] In one embodiment, the isolation devices 118 are distributed across the memory module 110 or the module board 119 in positions corresponding to the respective groups of

memory devices. For example, isolation device ID-1 is disposed in a first position corresponding to the first group of memory devices M_{i1} , M_{i2} , M_{i3} , and M_{i4} , and isolation device ID- i is disposed in an i^{th} position separate from the first position and corresponding to the i^{th} group of memory devices M_{i1} , M_{i2} , M_{i3} , and M_{i4} . In one embodiment, the first position is between the first group of memory devices and an edge 201 of the module board 119 where connections (not shown) to the data/strobe signal lines 130 are placed, and i^{th} position is between the i^{th} group of memory devices and the edge 201 of the module board 119. In one embodiment, the isolation devices 118 are distributed along the edge 201 of the memory module 110. In one embodiment, each isolation device 118 is a separate integrated circuit device packaged either by itself or together with at least some of the respective group of memory devices. In one embodiment, the module data/strobe signal lines 210, the module C/A signal lines 220, and the module control signal lines 230 include signal traces formed on and/or in the module board 119.

[0048] As an option, memory module 110 may further include a serial-presence detect (SPD) device 240, which may include electrically erasable programmable read-only memory (EEPROM) for storing data that characterize various attributes of the memory module 110. Examples of such data include a number of row addresses, a number of column addresses, a data width of the memory devices, a number of ranks on the memory module 110, a memory density per rank, a number of memory device on the memory module 110, and a memory density per memory device, etc. A basic input/output system (BIOS) of system 100 can be informed of these attributes of the memory module 110 by reading from the SPD 240 and can use such data to configure the MCH 101 properly for maximum reliability and performance.

[0049] In certain embodiments, the SPD 240 and/or the control circuit 116 store module configuration information, such as: memory space translation code, memory address mapping function code, input and output signals timing control information for the control circuit 116, input and output signals electrical and logical level control information for the control circuit 116, etc. In certain embodiments, the SPD 240 contains a system view of the module 110 which can be different from an actual physical construction of the module 110.

For example, the SPD 240 stores at least one memory operation parameter that is different from a corresponding memory operation parameter in a system memory controller setting. The SPD 240 may also store at least one data buffer operation parameter that is different from a corresponding parameter in the system memory controller setting.

[0050] Thus, in certain embodiments, in the memory module 110, C/A signals representing a memory command are received and buffered by the module control circuit 116, so that the MCH sees only the module control circuit 116 as far as the C/A signals are concerned. Write data and strobe signals from the controller are received and buffered by the isolation devices 118 before being transmitted to the memory devices 112 by the isolation devices 118. On the other hand, read data and strobe signals from the memory devices are received and buffered by the isolation devices before being transmitted to the MCH via the system data/strobe signal lines 130. Thus, MCH 101 does not directly operate or control the memory devices 112. As far as data/strobe signals are concerned, the MCH 101 mainly sees the isolation devices 118, and the system 100 depends on the isolation devices 118 to properly time the transmission of the read data and strobe signals to the MCH 101.

[0051] In certain embodiments, the memory module 110 is a dual in-line memory module (DIMM) and the memory devices are double data rate (DDR) dynamic random access memory devices (DRAM). In certain embodiments, the control circuit 116 includes a DDR register, and logic for memory space translation between a system memory domain and a module level physical memory domain. Such translation may produce address mapping, proper interface timing for the control signals to the module level physical memory domain, and a proper interface electrical and logical level for the control signals to the module level physical memory domain.

[0052] As shown in FIG. 2C, in certain embodiments, the control circuit 116 transmits registered C/A and clock signals to the memory devices 112, and transmits module control signals and a registered clock signal (or module clock signal) to the isolation devices 118, in a fly-by configuration. As the speed of memory operations increase, issues can arise with respect to signal alignment for input, output delay variation due process, voltage and

temperature (PVT) variations, synchronization with system memory controller interface, and phase drift accumulation during operation, etc. Electrical interface calibration drift during operation due to charge build up and timing interface calibration drift during operation due to environment change can also create issues.

[0053] For example, load reduction mechanism in the isolation devices 118 would provide a single data bus interface to the respective set of memory devices, which is hidden from the system memory controller 101. Thus, a long sequence of interface timing training may be required due to limited controllability of the system memory controller 101 over the interface between the memory devices 112 and the associated isolation devices 118. Furthermore, interface signal alignment-drift after the initial training would not be easily detected by the system memory controller 101, which may cause silent system failure.

[0054] Moreover, clock skew amongst the memory devices 112 and the associated isolation devices 118 due to the distributed architecture of the memory module 110 can cause synchronization issues. As the speed of memory operation increase, data period can become very close to the signal propagation delay time. Thus, such issues cannot simply be addressed by pipelining the data paths, as variation of the signal propagation time through I/Os becomes a very significant portion of a data period.

[0055] To address at least some of the above issues, in certain embodiments, as shown in FIG. 2D, the control circuit 116 transmits registered C/A signals to the memory devices 112, and transmits the module control signals and the module clock signal to the data buffers 118, in a fly-by arrangement. The memory devices 112 do not receive the module clock signal from the control circuit 116. Instead, each data buffer 118 regenerates the clock that is used by the respective set of memory devices 112. Each Data buffer 118 is thus responsible for providing a correct data timing interface between the respective set of memory devices 112 and the system memory controller 101. Each data buffer 118 is also responsible for providing the correct control signal timing between the control circuit 116 and the respective set of memory devices 112.

[0056] Thus, the memory module 110 in FIG. 2D allows a locally synchronized operation for each respective set of memory devices 112, which can correspond to a nibble or a byte of a DDR data bus between the memory module 110 and the system memory controller 101. Also, signal interface between each data buffer 118 and the respective set of memory devices 112 can be synchronized. In one embodiment, each data buffer 118 has a set of configurable operations, including, for example: programmable phase relationship between the clock it receives and the clock it regenerates, programmable phase adjustment for the data and data-strobe signals coupled to the memory devices 112, programmable phase adjustment for the data and data-strobe signals coupled to the system memory controller 101, programmable phase adjustment related to at least one control signal that is coupled to the control circuit 116. The locally synchronized operation also makes it easier for each data buffer 118 to perform self-testing of the associated set of memory devices 112, independent of the self-testing of other sets of memory devices performed by the other data buffers, as disclosed in commonly-owned U.S. Patent 8,001,434, entitled “Memory Board with Self-Testing Capability,” which is incorporated herein by reference in its entirety.

[0057] In certain embodiments, operations of the isolation devices 118 are controlled by the module control signals from the module control circuit 116, which generates the module control signals according to the C/A signals received from the MCH. Thus, the module control signals need to be properly received by the isolation devices 118 to insure their proper operation. In one embodiment, the module control signals are transmitted together with the module clock signal CK, which is also generated by the module control circuit 116 based on the system clock signal MCK. The isolation circuits 118 buffers the module clock signal, which is used to time the sampling of the module control signals. Since the isolation devices 118 are distributed across the memory module, the module control signal lines 230 can stretch across the memory module 110, over a distance of several centimeters. As the module control signals travel over such a distance, they can become misaligned with the module clock signal, resulting in metastability in the received module control signals. Therefore, in one embodiment, the isolation circuits 118 includes metastability detection circuits to detect metastability condition in the module control

signals and signal adjustment circuits to adjust the module control signals and/or the module clock signal to mitigate any metastability condition in the module control signals, as explained in further detail below.

[0058] Because the isolation devices 118 are distributed across the memory module 110, during high speed operations, it may take more than one clock cycle time of the system clock MCK for the module control signals to travel along the module control signals lines 230 from the module control device 116 to the farthest positioned isolation devices 118, such as isolation device ID-1 and isolation device ID-($n-1$) in the exemplary configuration shown in FIG. 2. In other words, a same set of module control signals may reach different isolation devices 118 at different times across more than one clock cycle of the system clock. For example, when the clock frequency of the system clock is higher than 800 MHz, the clock cycle time is less than about 1.2 ns. With a signal travel speed of about 70ps per centimeter of signal line, a module control signal would travel about 15 cm during one clock cycle. When the clock frequency increases to 1600 MHz, a module control signal would travel less than 8 cm during one clock cycle. Thus, a module control signal line can have multiple module control signals on the line at the same time, i.e., before one module control signal reaches an end of the signal line, another module control signal appear on the signal line.

[0059] With the isolation devices 118 receiving module control signals at different times across more than one clock cycle, the module control signals alone are not sufficient to time the transmission of read data signals to the MCH 101 from the isolation devices 118. In one embodiment, each isolation devices includes signal alignment circuits that determine, during a write operation, a time interval between a time when one or more module control signals are received from the module control circuit 116 and a time when a write strobe or write data signal is received from the MCH 101. This time interval is used during a subsequent read operation to time the transmission of read data to the MCH 101, such that the read data follows a read command by a read latency value associated with the system 100, as explained in more detail below.

[0060] More illustrative information will now be set forth regarding various optional configurations, architectures, and features with which the foregoing framework may or may not be implemented, per the desires of the user. It should be strongly noted that the following information is set forth for illustrative purposes and should not be construed as limiting in any manner. Any of the following features may be optionally incorporated with or without the exclusion of other features described.

[0061] In one embodiment, as shown in FIG. 3, each group of signal lines 130 include a set of n data (DQ) signal lines 322 each for transmitting one of a set of data signals $DQ_0, DQ_1, \dots, DQ_{n-1}$, and at least one strobe (DQS) signal line 324 for transmitting at least one strobe signal DQS. Each set of module data/strobe lines Y include a set of n module data signal lines Y_0, Y_1, \dots, Y_{n-1} and at least one module strobe signal line Y_{DQS} . When the subsets of memory devices are coupled to the associated isolation device 118 via respective subsets of memory devices, each set of module data/strobe lines Y may include multiple subsets of module data/strobe lines, such as the subsets of module data/strobe lines YA and YB shown in FIG. 2B. Each subset of module data/strobe lines YA include a set of n first module data lines YA_0, YA_1, \dots, YA_n and at least one first module strobe signal line YA_{DQS} ; and each subset of module data/strobe lines YB include a set of n second module data lines YB_0, YB_1, \dots, YB_n and at least one second module strobe signal line YB_{DQS} .

[0062] Each isolation device 118 includes a set of DQ routing circuits 320 coupled on one side to respective ones of the set of n DQ signal lines 322, and on another side to respective ones of the respective set of n module data lines, or respective ones of the respective subsets of module data lines, such as the first module data lines YA_0, YA_1, \dots, YA_n and the second module data lines YB_0, YB_1, \dots, YB_n . Each isolation device 118 further includes an ID control circuit 310 coupled on one side to the at least one DQS signal line 324, on another side to the one or more module strobe signal lines Y_{DQS} , or the first module strobe signal line YA_{DQS} and second module strobe signal line YB_{DQS} . The ID control circuit 310 also receives the module clock signal CK and the module control signals via the module control signal lines 230, and outputs ID control signals 330 to the DQ routing circuits 320, including, for example, one or more enable signals ENA and/or ENB , and some or all of

the other received, decoded, and/or otherwise processed module control signals, a delay signal DS, a read DQS signal RDQS, a write DQS signal WDQS, and a buffer clock signal CK0. Each DQ routing circuit 320 is configured to enable data communication between the respective DQ signal line 322 with a selected subgroup of one or more memory devices in response to the module control signals, as explained in more detail below.

[0063] In certain embodiments, the ID control circuit 310 also provides a delay signal DS, which is used by the DQ routing circuits 320 to align read data output by the isolation device 118 with read data output by the other isolation devices 118, as explained in further detail below. In certain embodiments, the ID control circuit 310 regenerates a clock signal from the module clock signal CK, which can have a programmable delay from the module clock signal. The regenerated clock signal is used as the clock signal CK0 and a clock signal CKM that is provided to the corresponding set of memory devices, as explained in more detail below.

[0064] The memory devices 112 are coupled to the isolation devices 118 via a same set of module data/strobe signal lines or different subsets of module data/strobe signal lines. For example, as shown in FIG. 4A, memory devices M_{11} , M_{12} , M_{13} , and M_{14} in the first group of memory devices can be coupled to the isolation device ID-1 via a same set of module data lines $Y-1_0$, $Y-1_1$, ..., $Y-1_{n-1}$ and module strobe line $Y-1_{DQS}$. In such embodiment, a subgroup in the group of memory devices can be selected by the isolation devices to communicated data with the MCH based on the phases of the data/strobe signals, which can be different with respect to different subgroups of memory devices.

[0065] Alternatively, as shown in FIG. 4B, memory devices M_{11} and M_{13} , which form a subgroup in the first group of memory devices, are coupled to the isolation device ID-1 via the module data lines $YA-1_0$, $YA-1_1$, ..., $YA-1_n$ and module strobe line $YA-1_{DQS}$ and memory devices M_{12} and M_{14} , which form another subgroup in the first group of memory devices, are coupled to the isolation device ID-1 via the module data lines $YB-1_0$, $YB-1_1$, ..., $YB-1_n$ and module strobe line $YB-1_{DQS}$. Memory devices coupled to the same isolation devices can be disposed on a same side or different sides of the memory board

119. Memory devices coupled to the same isolation devices may be placed side by side, on opposite sides of the module boards 119, or stacked over each other, and/or over the associated isolation device.

[0066] Multiple memory devices having a data width that is less than a data width of the isolation devices 118 may be used in place of one of the memory devices 112, which has the same data width as that of the isolation devices. For example, as shown in FIG. 5A, two memory devices M_{11-1} and M_{11-2} may be used in place of the memory device M_{11} . Each of the two memory devices M_{11-1} and M_{11-2} has a data width of 4, and together they act like a memory device M_{11} of a data width of 8. Thus, memory device M_{11-1} is coupled to the isolation device ID-1 via module data lines $YA-1_0, \dots, YA-1_3$ and module strobe line $YA-1_{DQS-1}$ while memory circuit M_{11-2} is coupled to the isolation device ID-1 via module data lines $YA-1_4, \dots, YA-1_7$ and module strobe line $YA-1_{DQS-2}$.

[0067] In another embodiment, as shown in FIG. 5B, four memory devices M_{11-1} to M_{11-4} may be used as the memory device M_{11} . Each of the four memory devices M_{11-1} to M_{11-4} has a data width of 4, and together they act like a memory device M_{11} of a data width of 16. Thus, memory device M_{11-1} is coupled to the isolation device ID-1 via module data lines $YA-1_0, \dots, YA-1_3$ and module strobe line $YA-1_{DQS-1}$ while memory device M_{11-2} is coupled to the isolation device ID-1 via module data lines $YA-1_4, \dots, YA-1_7$ and module strobe line $YA-1_{DQS-2}$, and so on.

[0068] FIG. 6 illustrates the ID control circuit 310 in an isolation device 118. As shown, the ID control circuit 310 includes a clock buffer 610 to receive the module clock signal CK from the module control device 116, and to output a module clock signal CK0. The ID control circuit 310 further includes a strobe routing circuit 620 that are coupled on one side to the corresponding system DQS signal line 324 and on another side to the corresponding module DQS signal lines YA_{DQS} and YB_{DQS} . The ID control circuit 310 further includes a receiver circuit 630 with respect to each of at least some of the module control signals (MCS) to receive a respective one of the module control signals. The ID control circuit 310 further includes a command processing circuit 640 that provides the received, decoded,

and/or otherwise processed module control signals 330 to the DQ routing circuits 320 and the strobe routing circuit 620 either directly or after further processing, if needed. The received/decoded/processed module control signals may include, for example, one or more enable signals ENA and/or ENB that are used by the DQ routing circuits 320 and the strobe routing circuit 620 to selectively enabling data communication between the MCH 101 and one of the subgroups in the respective group of memory devices, with which the isolation device is associated.

[0069] The strobe routing circuit 620 also buffers strobe signals received from either the MCH 101 or the memory devices 112, and output either a write strobe WDQS or read strobe RDQS to the DQ routing circuits 320. In one embodiment, the ID control circuit 310 further includes a delay control circuit 650 that receives one of the module control signals and either a data signal or a strobe signal and determines a delay amount to be used by the DQ routing circuit 320 and the strobe routing circuit 620. The delay amount is provided to the DQ routing circuit 320 and the strobe routing circuit in a delay signal DS.

[0070] In a receiver circuit 630, the respective MCS is received in accordance with the module clock signal CK0. In one embodiment, receiver circuit 630 samples the respective MCS using rising (or falling) edges of the module clock CK0. Since the isolation devices 118 are distributed across the memory module 110 at positions corresponding to the respective groups of memory devices, the module control signal lines 230 that carry the MCS to the isolation devices can stretch over a distance of more than 10 centimeters, as shown in FIG. 7. As the MCS and CK0 travel along their respective module control signal lines 710 and 720, they can become misaligned with each other when they reach the input pins 730 of an isolation device 118.

[0071] For example, a module control signal, like the MCS 810 shown in FIG. 8, can be perfectly aligned with the module clock signal CK, with a rising edge 801 of the module clock signal CK being at a center of a data eye 802, when the MCS signal and the clock signal leave the module control circuit 116. When the module control signal and the module clock signal reach an isolation device, however, their alignment can become shifted

like the MCS 820 with respect to the CK signal, i.e., the rising edge 801 of the clock signal is near a left edge of a data eye of the MCS 820, barely providing enough set up time for proper sampling of the module control signal. Or, the module control signal, like the MCS 830, can be shifted with respect to the module clock signal such that a rising edge 801 of the clock signal is near a right edge of a data eye of the MCS, barely providing enough hold time for proper sampling of the module control signal. Or, ever worse, the module control signal, like the MCS 840, can be so shifted with respect to the module clock signal such that a rising edge 801 of the clock signal falls in the glitches 803 at the edge of a data eye of the MCS, meaning that the sampled results could be metastable.

[0072] In one embodiment, as shown in FIG. 9, a receiver circuit 630 includes a metastability detection circuit (MDC) 910 to determine a metastability condition in a corresponding module control signal MCS0. In one embodiment, the MDC 910 generates at least one delayed version of the module clock signal CK and at least one delayed version of the corresponding MCS0. The MDC 910 also generates one or more metastability indicators and outputs the one or more metastability indicators via lines 912 and/or 914.

[0073] The receiver circuit 630 further includes a signal selection circuit 920 that receives the module clock CK and the at least one delayed version of the module clock via signal lines 916. The signal selection circuit 920 also receives the corresponding MCS and the at least one delayed version of the corresponding MCS via signal lines 918. The signal selection circuit 920 selects a clock signal CK_i from among the module clock CK and the at least one delayed version of the module clock based on one or more of the metastability indicators. The signal selection circuit 920 may also select an MCS signal MCS_i from among the corresponding MCS and the at least one delayed version of the corresponding MCS based on at least one other metastability indicator.

[0074] The receiver circuit 630 further includes a sampler or register circuit 930 that samples the selected module control signal MCS_i according to the selected clock signal CK_i and outputs the sampled signal as the received module control signal, which is provided to the command processing circuit 640 for further processing (if needed) before being

provided to the DQ routing circuits 320 and DQS routing circuit 620.

[0075] FIG. 10A illustrates an MDC 910 according to one embodiment. As shown, the MDC 910 includes a delay circuit 1012 that generates a delayed version MCS1 of the corresponding MCS0 by adding a predetermined amount of delay (e.g., 10ps) to MCS0. MDC 910 also includes a delay circuit 1016 that generates a delayed version CK1 of the clock signal CK0 by adding a predetermined amount of delay to CK0. In one embodiment, CK1 is delayed from CK0 by about 1/10th of a clock cycle, e.g., 50-70 ps for an operating frequency of about 1600 MHz. The MDC 910 further includes a sampler circuit 1042 that samples MCS1 according to CK0 and outputs a sampled result A, a sampler circuit 1044 that samples MCS0 according to CK0 and outputs a sampled result B, and a sampler circuit 1046 that samples MCS0 according to CK1 and outputs a sampled result C. The MDC 910 further includes a logic circuit (e.g., a majority decision circuit) that generates metastability indicators Z1 and Z2 based on the sampled results A, B, and C.

[0076] In one embodiment, Z1 is the result of a logic operation (e.g., an XNOR operation) on the sampled result, e.g., $Z1 = \overline{A \oplus B}$, and Z2 is the result of another logic operation on the sampled results, e.g., $Z2 = \overline{B \oplus C}$. Thus, as shown in FIG. 10B and Table 1 below, when a metastability condition of insufficient hold time occurs, i.e., a rising clock edge 1061 of CK0 is close to the right side of a data eye where glitches at the edges of the data eyes can make C unpredictable, A and B can be in agreement (i.e., Z1 is true) while B and C are likely not in agreement (i.e., Z2 is false). FIG. 10 C illustrates a metastability condition when there is insufficient set-up time. As shown in FIG. 10C and Table 1 below, a rising clock edge 1061 of CK0 is close to the left side of a data eye where glitches at the edges of the data eyes can make A unpredictable. Thus, A and B can be in disagreement so Z1 is false while B and C can be in agreement so Z2 is true. Not shown in the figures is the situation that all A, B, and C are in agreement, meaning that both the rising clock edge 1061 of CK0 and the rising clock edge 1062 of CK1 are near the middle of an MCS0 data eye so there is no metastability issues and both Z1 and Z2 are true, as shown in Table 1.

[0077] FIG. 10D illustrates a signal selection circuit 920 according to an embodiment. As

shown, in one embodiment, the signal selection circuit 920 includes a first multiplexor 1071 that selects between CK0 and CK1 based on the metastability indicator Z1, and a second multiplexor 1072 that selects between MCS0 and MCS1 based on the metastability indicator Z2. Thus, as shown in Table 1, where a metastability condition of insufficient hold time occurs, Z1=1 and Z2=0, and MCS1 is output from multiplexor 1071 while CK0 is output from multiplexor 1072. Sampler 930 thus samples MCS1 according to the rising edges of CK0. Thus, more hold time is provided to mitigate the metastability condition since MCS1 is shifted from MCS0 toward the right.

[0078] On the other hand, where a metastability condition of insufficient set-up time occurs, Z1=0 and Z2=1, and CK1 is output from multiplexor 1071 while MCS0 is output from multiplexor 1072. Sampler 930 thus samples MCS0 according to the rising edges of CK1. Since CK1 is shifted from CK0 toward the right, more set-up time is provided to mitigate the metastability condition.

[0079] In the case when no metastability is detected, Z1=1 and Z2=1, and CK0 is output from multiplexor 1071 while MCS0 is output from multiplexor 1072. So, the unshifted module control signal is sampled according to the unshifted module clock signal.

Table 1 Metastability Detection and Signal Selection

Sampler Output			MS Indicators		MS Condition	Signal Selection	
A	B	C	Z1	Z2		CK	MCS
D1	D1	D2	1	0	insufficient hold time	CK0	MCS1
D1	D2	D2	0	1	insufficient set-up time	CK1	MCS0
D1	D1	D1	1	1	no metastability	CK0	MCS0

[0080] FIGS. 10A-10D illustrate a relatively simple implementation of the metastability detection circuit (MDC) 910 where only three different sample points are provided to detect metastability condition in the module control signal. In general, the MDC 910 may generate more delayed versions of the module clock signal CK0 and/or the corresponding module control signal MCS0, and may include more sampler circuits to sample any

additional delayed versions of the module control signal according to either the module clock signal or one of the delayed versions of the module clock signal. For example, as shown in FIG. 11A, the MDC 910 can include a plurality of delay circuits 1102 that generate m delayed versions of MCS0, e.g., MCS1, MCS2, ... MCS m , and m delayed versions of CK0, e.g., CK1, CK2, ..., CK m . The MDC 910 can include sampler circuits 1104 that sample MCS0 according to CK0, CK1, ..., CK m , respectively, and sampler circuits 1104 that sample MCS0, MCS1, MCS2, ... MCS m according to CK0, respectively. The outputs of the samplers 1104 are provided to a logic circuit 1120, which determines a metastability condition in MCK0 based on the sampler outputs using, for example, a majority decision logic. The logic circuit 1120 outputs a first metastability indicator on line(s) 912 and a second metastability indicator on line(s) 914.

[0081] FIG. 11B illustrates a signal selection circuit 920 according to an embodiment. As shown, in one embodiment, the signal selection circuit 920 includes a first multiplexor 1171 that selects between CK0, CK1, ..., CK m based on the metastability indicator provided on line(s) 912, and a second multiplexor 1172 that selects between MCS0, MCS1, ..., MCS m based on the metastability indicator provided on line(s) 914, such that the rising edges of the selected clock signal, e.g., Ck i , are close to the middle of the respective data eyes in the selected module control signal, e.g., MCS i . The selected signals MCS i and Ck i are provided to the sampler 930, which samples MCS i according to the rising edges of Ck i .

[0082] As stated above, in certain embodiments, since the isolation devices 118 are in the data paths between the MCH 101 and the respective groups of memory devices 112, the MCH 101 does not have direct control of the memory devices 112. Thus, conventional read/write leveling techniques are not sufficient for managing read/write data timing. In one embodiment, the isolation devices 118 includes signal alignment mechanism to time the transmission of read data signals based on timing information derived from a prior write operation, as discussed further below.

[0083] FIG. 12A is a timing diagram for a write operation according to one embodiment.

As shown, after a write command W/C associated with the write operation is received by the module control circuit 116 at time t1, the module control circuit 116 outputs one or more enable signals EN at time t2 in response to the write commands. The one or more enable signals are received by an isolation device 118 at time t3, which afterwards receives one or more strobe signal DQS from the MCH 101 at time t4. Note that the same enable signal may be received by another isolation device 118 at time t3', which can be in a different cycle of the system clock MCK from the cycle which t3 is in. The time interval between t4 and t1 is consistent with a write latency W.L. associated with the system 100, and is controllable by the MCH 101 and knowable to the isolation device 118. The time interval between t4 and t3, referred to hereafter as an enable-to-write data delay EWD, can be determined by the isolation device 118 since both these signals are received by the isolation device. Based on such determination, the isolation device 118 can have knowledge of the time interval between t3 and t1, referred to hereafter as a command-to-enable delay CED, which can be used by the isolation device 118 to properly time transmission of read data to the MCH, as explained further below.

[0084] FIG. 12B is a timing diagram for a read operation according to one embodiment. As shown, after a read command R/C associated with the read operation is received by the module control circuit 116 at time t5, the module control circuit 116 outputs one or more enable signals EN at time t6 in response to the read commands. The one or more enable signals are received by an isolation device 118 at time t7, which afterwards receives at time t8 read data signals (not shown) and one or more strobe signal DQS from the respective group of memory devices. Note that the same enable signal may be received by another isolation device 118 at time t3', which can be in a different cycle of the system clock MCK from the cycle which t3 is in. Thus, the enable signals alone cannot be used to time the transmission of the read signals by the isolation devices 118.

[0085] With knowledge of the time interval between t7 and t5, which should be about the same as the time interval between t3 and t1, i.e., the command-to-enable delay CED, in certain embodiments, the isolation device can add a proper amount of delay to the read data signals and the one or more DQS signal such that the read data signals and the one or more

DQS signal are transmitted at time t_9 by the isolation device to the MCH 101 via the respective group of data/strobe signal lines 130, with the time interval between t_9 and t_5 being consistent with a read latency R.L. associated with the system 100.

[0086] The time interval between t_4 and t_3 , i.e., the enable to write data delay EWD, is determined by the delay control circuit 650 in the ID control circuit 310, as shown in FIG. 6. According to one embodiment, as shown in FIG. 13, the delay control circuit 650 includes a preamble detector 1310 to detect a write preamble in the DQS, a flip-flop circuit 1320 having an enable input EN receiving one of the module control signals and a clock input CK receiving the buffered module clock signal CK0, and a counter circuit 1330 having a Start input receiving the one of the module control signals, a Stop input receiving an output of the flip-flop circuit 1320. Thus, the output of the counter circuit, i.e., the delay signal DS, would indicate a time interval from when the write preamble is detected and when the one of the module control signal is received.

[0087] FIG. 14 illustrates a DQ or DQS routing circuit 320 or 620 according to an embodiment. As shown, the DQ/DQS routing circuit 320/620 includes a DQ/DQS pin 1401 that is coupled to the corresponding DQ/DQS signal line 322/324, a set of one or more DQS pins 1402 that is coupled to a corresponding module DQ/DQS line(s) Y/Y_{DQS}, or YA/YA_{DQS} and YB/YB_{DQS}. The DQ/DQS routing circuit 320/620 further includes a write strobe buffer 1410 that buffers write data/strobe, and a write data/strobe receiver 1420 that samples the write data/strobe. The DQ/DQS routing circuit 320/620 further includes a plurality of write paths 1430 that are selectable or can be selectively enabled by one or more of the module control signals, such as the enable signals ENA and ENB.

[0088] The DQS routing circuit further includes a plurality of read paths 1450 that are selectable by the one or more of the module control signals. Output from the selected read path is delayed in a delay circuit 1460 by an amount controlled by the delay signal DS, and sampled by a sampler circuit 1470. The sampled read data/strobe is transmitted by transmitter 1480 onto the corresponding data/strobe signal line 322/324 via the DQ/DQS pin 1401.

[0089] FIG. 15 illustrates a DQS routing circuit 620 according to an embodiment. As shown, the DQS routing circuit 620 includes a first DQS pin 1501 that is coupled to a corresponding DQS signal line 324, a second DQS pin 1502A that is coupled to a corresponding module DQS line Y_{ADQS} , a third DQS pin 1502B that is coupled to a corresponding module DQS line Y_{BDQS} . The DQS routing circuit 620 further includes a first write strobe path coupled between the first DQS pin 1501 and the second DQS pin 1502A and a second write strobe path coupled between the first DQS pin 1501 and the third DQS pin 1502B. The first write strobe path includes a write strobe buffer 1510 that buffers a write strobe, a write strobe receiver 1520 that samples the write strobe according to the buffered module signal CK0. The sampled write strobe is provided to the DQ routing circuits 320 as the write strobe WDQS. The first write strobe path further includes a first write strobe transmitter 1530A that transmits the write strobe to one or more memory devices 112 coupled to the module strobe line Y_{ADQS} . The second write strobe path includes the write strobe buffer 1510, the write strobe receiver 1520, and a second write strobe transmitter 1530B that transmits the write strobe to one or more memory devices 112 coupled to the module strobe line Y_{BDQS} . The first and second write strobe transmitters, 1530A and 1530B, are controlled by two enable signals, ENA and ENB, respectively, such that the first write strobe path and the second write strobe path can be selectively enabled/disabled by the enable signals, ENA and ENB.

[0090] The DQS routing circuit further includes a read strobe path coupled between the first DQS pin 1501 and a selected one of the second and third DQS pins 1502A and 1502B. In the read strobe path, a select circuit 1550 (e.g., a multiplexor) selects either a read strobe signal received via DQS pin 1502A or a read strobe signal received via DQS pin 1502B based on one or both of the enable signals ENA or ENB. The selected read strobe signal is delayed in a delay circuit 1560 by an amount controlled by the delay signal DS, and sampled by a sampler circuit 1570 according to the buffered module clock signal CK0. The sampled read strobe is provided to the DQ routing circuits 320 as the read strobe RDQS and is transmitted by transmitter 1580 onto the corresponding strobe signal line 324 via the first DQS pin 1501.

[0091] FIG. 16 illustrates a DQ routing circuit 320 according to an embodiment. As shown, the DQ routing circuit 320 includes a first DQ pin 1601 that is coupled to a corresponding DQ signal line 130, a second DQ pin 1602A that is coupled to a corresponding module DQ line YA_{DQ} , a third DQ pin 1602B that is coupled to a corresponding module DQ line YB_{DQ} . The DQ routing circuit 320 further includes a first write data path coupled between the first DQ pin 1601 and the second DQ pin 1602A and a second write data path coupled between the first DQ pin 1601 and the third DQ pin 1602B. The first write data path includes a write data buffer 1610, a write data receiver 1620 that samples write data according to the write strobe $WDQS$ from the DQS routing circuit 620, and a first write data transmitter 1630A that transmits the write data to one or more memory devices 112 coupled to the module data line YA_{DQ} . The second write data path includes the write data buffer 1610, the write data receiver 1620, and a second write data transmitter 1630B that transmits the write data to one or more memory devices 112 coupled to the module data line YB_{DQ} . The first and second write data transmitters, 1630A and 1630B, are controlled by two enable signals, ENA and ENB , respectively. Thus, the first write data path and the second write data path can be selectively enabled/disabled by the enable signals, ENA and ENB .

[0092] The DQ routing circuit further includes a read data path coupled between the first DQ pin 1601 and a selected one of the second and third DQ pins 1602A and 1602B. In the read data path, a select circuit 1650 (e.g., a multiplexor) selects either a read data signal received via DQ pin 1602A or a read data signal received via DQ pin 1602B based on one or both of the enable signals ENA or ENB . The selected read data signal is delayed in a delay circuit 1660 by an amount controlled by the delay signal DS . The delayed read data signal is then sampled by a receiver circuit 1670 according to the read strobe $RDQS$ from the DQS routing circuit 620, and transmitted by transmitter 1680 onto the corresponding data signal line 130 via the first DQ pin 1601.

[0093] FIG. 17 illustrate a delay circuit 1560 or 1660 according to an embodiment. As shown, the delay circuit 1560 or 1660 includes a plurality of delay stages, such as delay stages 1710, 1720, and 1730, each delaying a read data or read strobe signal from the select

circuit 1550/1650 by a predetermined amount. The delay circuit 1560 or 1660 further includes a select circuit 1740 (e.g., a multiplexor) that selects from among the read data or read strobe signal and the outputs from the delay stages according to the delay signal DS. The output of the select circuit 1740, is provided to the sampler circuit 1570 or 1670, either directly or after being buffered by a buffer circuit 1750.

[0094] Thus, as shown in FIG. 18, in one embodiment, a memory module 110 operates in the memory system 100 according to a method 1800. In the method, during a write operation, one or more module control signals are received by an isolation device 118 from a module control circuit or module controller 116 (1810). The module controller 116 generates the one or more module control signals in response to C/A signals representing a write command from the MCH 101. The one or more module control signals are used to control the isolation device 118. For example, the one or more module control signals may include one or more first enable signals to enable a write path to allow write data be communicated to a selected subgroup of memory devices among the group of memory devices coupled to the isolation device 118. After a time interval from receiving the one or more first enable signals, write data DQ and write strobe DQS are received by the isolation device 118 from the MCH 101 (1820). In one embodiment, upon receiving the one or more first enable signal, a counter is started, which is stopped when the write data DQ or write strobe DQS is received. Thus, a time interval EWD between receiving the one or more first enable signals and receiving the write strobe signal DQS is recorded.

[0095] Since the time interval between the arrival of the command signals from the MCH 101 and the arrival of the write data/strobe signal DQ/DQS from the MCH 101 is a set according to a write latency parameter associated with the system 100, the time interval EWD can be used to ascertain a time interval CED between the time when a command signal is received by the memory module 110 and the time when the one or more enable signals are received by the isolation device 118. The time interval CED can be used by the isolation device 118 to properly time the transmission of read data to the MCH 101, as described above and explained further below.

[0096] As shown in FIG. 18, a delay signal DS is generated according to the time interval EWD (1830). Concurrent to receiving the write strobe signal DQS, the isolation device 118 also receives a set of write data signals DQ (1840). The received write data signals are transmitted to the subgroup of memory devices (1850), which are selected from the group of memory devices coupled to the isolation device 118 by the one or more first enable signals.

[0097] During a read operation, another set of module control signals including, for example, one or more second enable signals, are received by the isolation device 118 from the module controller 116 (1860). The one or more second enable signals are generated by the module controller 116 in response to read command signals received from the MCH 101, and are used by the isolation device 118 to select a subgroup of memory devices from which to receive read data. Afterwards, a read strobe signal DQS and a set of read data signal DQ are received from the selected subgroup of memory devices (1870). To properly time the transmission of the DQS and DQ signals to the MCH 101, the DQS and DQ signals are adjusted (e.g., delayed) according to the delay signal DS, such that the DQS and DQ signals follow a read command by a time interval consistent with a read latency parameter associated with the system 100.

[0098] In certain embodiments, especially the embodiments shown in FIG. 2D, the delay circuits 1560 and 1660 shown in FIGS. 15 and 16 are not needed to provide alignment of the read data. As shown in FIG. 19, the ID control circuit 310 includes a clock regeneration circuit 1920 that regenerates the clock signal CK received from the control circuit 116, according to the delay signal DS. The regenerated clock signals CK0 and CKM each includes a proper amount of delay as compared to the clock signal CK. The clock CK0 is provided to the strobe routing circuit 620 so that the strobe signals are properly timed to result in proper data alignment. The regenerated clock signal CKM is provided to the respective set of memory devices so that the respective data buffer 118 and the respective set of memory devices are locally synchronized.

WE CLAIM:

1. A memory module to operate in a memory system with a memory controller, the memory system operating according to a system clock, comprising:
 - a module control device to receive command signals from the memory controller and to output module command signals and module control signals;
 - memory devices organized in groups, each group including at least one memory device, the memory devices receiving the module command signals from the module control device; and
 - a plurality of buffer circuits to receive the module control signals, a respective buffer circuit corresponding to a respective group of memory devices, the respective buffer circuit including data paths for communicating data between the memory controller and the respective group of memory devices, the data paths being controlled by at least one of the module control signals; andwherein the plurality of buffer circuits are distributed across a surface of the memory module such that each module control signal arrives at the plurality of buffer circuits at different points in time, and
- wherein the plurality of buffer circuits are configured to align read data signals received from the respective groups of memory devices such that the read data signals are transmitted to the memory controller from the memory module substantially aligned with each other and in accordance with a read latency parameter of the memory system.
2. The memory module of claim 1, wherein each module control signal arrives at the plurality of buffer circuits at different points in time across more than one clock cycle of the system clock, and wherein the plurality of buffer circuits are configured to align the read data signals such that they are transmitted by the buffer circuits to the memory controller within one system clock cycle.
3. The memory module of claim 1, wherein each buffer circuit includes receiver circuits to receive corresponding ones of the module control signals, each receiver circuit including a metastability detection circuit to detect metastability condition in the

corresponding module control signal, and wherein each buffer circuit further includes at least one signal adjustment circuit to adjust one or more the module control signals to mitigate any metastability condition in the module control signals.

4. The memory module of claim 1, wherein the module control device further receives the system clock from the system memory controller and transmits a registered clock signal to the plurality of buffer circuits, and wherein each respective buffer circuit regenerates one or more clock signals from the registered clock signal and provides a regenerated clock signal to the respective group of memory devices.

5. The memory module of claim 1, wherein each buffer circuit comprises:
a time interval determination circuit to receive, during a write operation, a first signal from the module controller and a second signal from the memory controller and to generate a delay signal indicating a time interval between the first signal and the second signal;

a delay circuit to delay a signal in the buffer circuit according to the delay signal indicating the time interval, the signal in the buffer circuit being one of a read data signal, a read strobe signal, and a regenerated clock signal.

6. A method of operating a memory module in a memory system having a memory controller, comprising:

receiving a read command from the memory controller;

generating module command signals and module control signals in response to the read command;

producing read data from the memory devices in response to the module command signals;

controlling a plurality of data paths in a plurality of data buffers using the module control signals; and

outputting read data from the plurality of data paths to the memory controller, wherein at least one signal in each of the plurality of data buffers is delayed such that the read data output from the plurality of data paths are aligned with each other and in

accordance with a read latency parameter of the memory system.

7. The method of claim 6, further comprising:
 - determining a metastability condition associated with a respective module control signal at each of the plurality of data buffers; and
 - adjusting the respective module control signal based on the metastability condition.
8. The method of claim 6, further comprising, in the each data buffer:
 - receiving at least one first module control signal generated in the memory module in response to a write command;
 - determining a time interval between receiving the at least one first module control signal and receiving a write signal associated with the write command from the memory controller, the write signal being one of a write data signal and a write strobe signal; and
 - delaying the at least one signal in the respective data buffer according to the time interval.
9. The method of claim 8 wherein the at least one signal is a clock signal.
10. The method of claim 8 wherein the at least one signal is at least one of a read data signal and a read strobe signal.
11. The method of claim 8, wherein the write signal is a write strobe signal, the method further comprising:
 - detecting a preamble of the write strobe signal to determine the time interval.
12. A buffer circuit for use in a memory module coupled to a memory controller, the memory module including a module controller, comprising:
 - a time interval determination circuit to receive, during a write operation, a first signal from the module controller and a second signal from the memory controller and to generate a delay signal indicating a time interval between the first signal and the second signal; and
 - a delay circuit to delay at least one signal in the buffer circuit according to the delay

signal indicating the time interval.

13. The buffer circuit of claim 12 wherein the at least one signal is a clock signal.
14. The buffer circuit of claim 12 wherein the at least one signal is at least one of a read data signal and a read strobe signal.
15. The buffer circuit of claim 12, wherein the first signal is an enable signal to enable a write data path in the buffer circuit.
16. The buffer circuit of claim 15, wherein the second signal is a write strobe signal.
17. The buffer circuit of claim 15, wherein the second signal has a preamble, and wherein the time interval determination circuit includes a preamble detector to detect the preamble of the second signal.
18. The buffer circuit of claim 18, wherein the time interval determination circuit further includes a counter to count a number of clock cycles between receipt of the first signal and detection of the preamble of the second signal.
19. The buffer circuit of claim 12, further comprising:
 - a clock regeneration circuit to receive a clock signal from the module controller and to regenerate a regenerated clock signal according to the delay signal indicating the time interval.
20. The buffer circuit of claim 12, wherein the buffer circuit is coupled to a respective set of memory devices on the memory module, and wherein the buffer circuit further comprises a clock regeneration circuit to receive a clock signal from the module controller and to provide a regenerated clock signal to the respective memory devices.

MEMORY MODULE WITH DISTRIBUTED DATA BUFFERS AND METHOD OF OPERATION

ABSTRACT OF THE DISCLOSURE

A memory module is operatable in a memory system with a memory controller. The memory module comprises a module control device to receive command signals from the memory controller and to output module command signals and module control signals. The module command signals are provided to memory devices organized in groups, each group including at least one memory device, while the module control signals are provided to a plurality of buffer circuits to control data paths in the buffer circuits. The plurality of buffer circuits are associated with respective groups of memory devices and are distributed across a surface of the memory module such that each module control signal arrives at the plurality of buffer circuits at different points in time. The plurality of buffer circuits are configured to align read data signals received from the memory devices such that the read data signals are transmitted to the memory controller from the memory module substantially aligned with each other and in accordance with a read latency parameter of the memory system.